

# Bad attitudes

Deniz Özyıldız (Konstanz) & Tom Roberts (ILLC, Amsterdam)

{deniz-ozyildiz@uni-konstanz.de, t.roberts@uva.nl}

September 22, 2023 @ TbiLLC — Handout at <https://rotom.github.io/badattitudes.pdf>

## 1 Introduction

How can we explain cross-linguistic patterns in the meaning of attitude predicates?

There are many **predicates that imply the truth or the falsity of their complement**:

Factives presuppose truth. Veridicals/anti-veridicals (‘be right/be wrong’) entail truth and falsity.

### (1) Factives

- a. Alice knows that it’s raining in Leiden.  $\rightsquigarrow$  It’s raining in Leiden.
- b. Alice doesn’t know that it’s raining in Leiden.  $\rightsquigarrow$  It’s raining in Leiden.

### (2) Anti-veridicals

- a. Alice is wrong that it’s raining in Leiden.  $\rightarrow$  It’s not raining in Leiden.
- b. Alice isn’t wrong that it’s raining in Leiden.  $\rightarrow$  It’s raining in Leiden.

A puzzling gap is that *contrafactuals*, which would presuppose that their complement is false, **are not attested** (Holton 2017, see also Strohmaier & Wimmer 2022).<sup>1</sup>

### (3) Contrafactuals (unattested)

- a. Alice shknows that it’s raining in Leiden.  $\rightsquigarrow$  It’s not raining in Leiden.
- b. Alice doesn’t shknow that it’s raining in Leiden.  $\rightsquigarrow$  It’s not raining in Leiden.

This is surprising because it’s very natural to ascribe false belief to people, using various means.

### (4) My colleagues believe that I’m in Konstanz.

$\rightsquigarrow$  I’m not in Konstanz. (Percus 2006; Sauerland 2007; Chemla 2008)

And because predicates that presuppose truth exist, so do ones that *entail* falsity.

| The contrafactive gap |                    |   |
|-----------------------|--------------------|---|
|                       | inference of truth | inference of falsity  |
| via entailment        | <i>be right</i>    | <i>be wrong</i>   |
| via presupposition    | <i>know</i>        | <div style="border: 1px dashed black; width: 100px; height: 20px;"></div> |

**Four criteria for contrafactivity**, making *shknow* a contrafactive counterpart of a “factive mental state verb” (Williamson 2000):

For helpful comments and suggestions, we would like to thank the participants of the MECORE workshop in Edinburgh, the MLC seminar in Amsterdam, and SuB 28 in Bochum, as well as Pranav Anand, Ryan Bochnak, Milica Denić, Kajsa Djärv, Emily Hanink, Dan Hoek, Alda Mari, Dean McHugh, Prerna Nadathur, Doris Penka, Jed Pizarro-Guevara, Sheng-Fu Wang, and Simon Wimmer. This work is in part made possible through the project “MECORE: Cross-linguistic investigation of meaning-driven combinatorial restrictions in clausal embedding” jointly awarded by the AHRC (AH/V002716) and DFG (RO 4247/5-1) to Romero, Uegaki and Roelofsen.

<sup>1</sup>See Anand & Hacquard (2014) for the argument that if a doxastic predicate triggers the inference that its complement is true or false, that inference is backgrounded: Presupposition is our criterion for contra-factivity, not entailment.

- (a) stative
- (b) monomorphemic
- (c) doxastic (i.e., indicates the subject believes its complement)
- (d) contrafactive (i.e., presupposes the negation of its complement)

Holton (2017) proposes an ontological source for the contrafactive ban:

- *That*-clause complements of factive verbs refer to facts.
- By analogy, if there are contrafactives, their *that*-clause complements must refer to contrafacts (= “facts, only false”).
- There are no contrafacts, hence no contrafactives.

However:

- This explanation is either unfalsifiable or just a restatement of the generalization<sup>2</sup>
- If correct, we then wonder why contrafacts are not part of our (linguistic) ontology anyway

**Today:** We concur with Holton’s generalization, but argue it follows from a more general constraint on lexicalization:

- (5) **The lexicalization constraint (informal)**<sup>3</sup> Satisfying the presupposition of a lexical item *L* must ‘lead to’ the at-issue content expressed by *L* (cf. Roberts & Simons 2022)

## 2 No known false belief verb presupposes its complement

Qua claim of non-existence, establishing the claim that there are no contra-factives requires requires checking every attitude predicate, or providing a proof or a reason for non-existence.

Following Strohmaier & Wimmer (2022), we examine a number of prominent candidates for contrafactives, but will argue that none of them fit the bill entirely.

**A (non-exhaustive) list of false belief verbs** Many attitude predicates have been described as giving rise to False Belief Inferences (FBIs) stronger than those associated with ‘believe’ or ‘think.’

- |  |                      |
|--|----------------------|
| • Mandarin <i>yǐwéi</i>  | Glass (2022)         |
| • Kipsigis <i>par</i>  | Bossi (2023)         |
| • Tagalog <i>akala</i>   | Kierstead (2013)     |
| • Turkish <i>zannet-</i> and <i>san-</i>   | Albayrak (2019)      |
| • Spanish <i>creerse</i> (believe-REFL), French <i>s’imaginer</i> (REFL-imagine) | Anvari et al. (2019) |

<sup>2</sup>Negative events and individuals *are* used to model semantic phenomena (Bernard & Champollion 2018; Bledin 2022), e.g., *I saw Mary not leave*. This makes it less obvious that contrafacts do not exist/are not useful methodologically.

<sup>3</sup>We focus on the presuppositions of attitude predicates, and leave other presupposition triggers for further research.

- Taiwanese Southern Min *liah8-tsun2* (estimate-accurate) Hsiao (2017)
- Dutch *wijsmaken*, German *weismachen* (wise-make) Hoeksema (2021)
- German *wähnen*, Dutch *wanen* Frege (1892); Sander (ms.)
- English *delude* (oneself), *fool*

We argue that on closer inspection, none of these predicates are true contrafactuals—either they are not stative or the FBI disappears in some contexts.<sup>4</sup>

### Describe a change of state (*wijsmaken* (also *delude*, *fool*, etc.)

Some predicates, e.g. Dutch *wijsmaken*, do seem like they might presuppose the falsity of their complement:

- (6) Ze **maakte** me **wijs** dat ze rijk was.  
 she made me wise that she rich was  
 ‘She fooled me into believing that she was rich.’ (Hoeksema 2021, ex. 3a)

But verbs of this kind (describing deception) lexicalize a change of state, namely, from not believing *p* to believing *p*.<sup>5</sup>

### FBI not always present (*wähnen*, *wanen*, *liah8-tsun2*)

German *wähnen* was discussed by Frege (1892) in contrafactive terms:

- (7) Bebel wähnt, daß durch die Rückgabe Elsaß-Lothringens Frankreichs  
 Bebel WÄHNEN that through the return Alsace-Lorraine.GEN France.GEN  
 Rachegelüste beschwichtigt werden können.  
 desire for revenge appeased become can  
 ‘Bebel fancies that the return of Alsace-Lorraine would appease France’s desire for revenge.’  
 a. Bebel believes that the return of Alsace-Lorraine would appease France’s desire for revenge.  
 b. the return of Alsace-Lorraine would not appease France’s desire for revenge.

But, judgments regarding the strength of *wähnen*’s FBI and whether it is always present are unclear and vary by speaker (Sander ms., Doris Penka, p.c.).

In some contexts, FBIs with *wähnen* seem marginal or nonexistent:

- (8) Ich **wähne** mich glücklich.  
 I WÄHNEN myself happy  
 ‘I consider myself happy.’  
 ✗ I am not happy. (Doris Penka, p.c.)

<sup>4</sup>A number of these predicates (*wijsmaken*, *liah8-tsun2*, *creerse*) are polymorphemic, so we might consider ruling them out on these grounds. However, the intent of the monomorphemicity criterion is to rule out verbs in which the division of semantic labor between falsity and belief is clearly delineated between morphemes (e.g. *mis-VERB*). The predicates we consider don’t obviously behave this way, so we will not reject them for being polymorphemic alone.

<sup>5</sup>Interestingly, the attitude holder of verbs of this kind is also not the subject.

This seems to suggest the FBI in other *wähnen* cases comes from a source other than contrafactivity.

### FBI cancellable/reinforceable (*yǐwéi*, *akala*, *par*)

These predicates imply false belief in some contexts (9), but the FBI is cancellable (10-11) or reinforceable (12) (Glass 2022, Bossi 2023, Jed Pizarro-Guevara, p.c., pace Kierstead 2013)<sup>6</sup>

- (9) tā **yǐwéi** tā huì yíng  
3s yǐwéi 3s will win  
‘She is under the impression that she’s going to win.’ (Glass 2022, ex. 5)  
↪ She isn’t going to win.
- (10) rénmen **yǐwéi** tā shì yìwànfùwēng... ér tā díquè shì.  
person-PL yǐwéi 3s be billionaire and 3s indeed be  
‘People are under the impression that she’s a billionaire... and she actually is.’ (Glass 2022, ex. 9)
- (11) Di ko alam kung nakascore siya, peru **akala** niya siguro nakascore siya.  
NEG I know Q was.able.score 3sg, but AKALA 3sg perhaps was.able.score 3sg  
‘I don’t know whether he was able to score, but he thinks that he might have scored.’ (Jed Pizarro-Guevara, p.c.)
- (12) **ø-par-e** kaameε-nyuun aa-mnyon-i lakini maa-mnyon-i.  
3-PAR-IPFV mother-my 1sg-be.sick-IPFV but NEG.1sg-be.sick-IPFV  
‘My mother is under the impression that I’m sick, but I’m not sick.’ (Bossi 2023, ex. 45)

⇒ The predicates *yǐwéi*, *akala* and *par* are not contrafactive because their FBI can be cancelled.

### FBI disappears under negation (*creerse*, *san-*)

The FBI associated with Spanish *creerse* is not cancelable, in (13), and Anvari et al. (2019) do analyze the predicate as presupposing the falsity of the complement.

- (13) Juan **se cree** que Ana tiene 30 años... #¡y tiene razón!  
Juan REFL believe that Ana has 30 years and has right  
Intended: Juan believes that Ana is 30 years old... and he’s right!  
(Anvari et al. 2019, ex. 10)

But when *creerse* is negated, it is reported to become factive, implying *p* rather than  $\neg p$ :

- (14) Juan no **se cree** que está lloviendo.  
Juan NEG REF believe that is raining  
**Available:**  
Juan doesn’t believe that it’s raining. (In the ‘can’t believe’ sense, Roberts 2019, 2021)  
↪ It’s raining. (Anvari et al. 2019, ex. 4)  
**Expected but unavailable:**  
Juan doesn’t believe that it’s raining.  
↪ It’s not raining.

<sup>6</sup>This is in part why Glass and Bossi do not argue that *yǐwéi* and *par* are contrafactive.

⇒ In that its FBI doesn't seem to project from under the scope of negation, *creerse* doesn't appear to presuppose falsity, and is not contrafactive either.

### Interim conclusion

- No predicate that we know of is contrafactive.
- If one is not convinced of the absence of contrafactives, they are at least rare (see Appendix) or their existence/behavior is less clear than that of regular factives.

## 3 Analysis

**Core intuition:** Contrafactivity is weird because the presupposition  $\neg p$  and the belief that  $p$  do not seem to have anything to do with one another  $\leadsto$  they don't belong in a single lexeme

Rather, it seems reasonable that lexically-encoded presuppositions of predicates need to somehow 'support' or 'lead to' the at-issue content of those predicates<sup>7</sup>

This can be seen as a more general version of the 'ontological precondition' view of presuppositions (Roberts & Simons 2022).<sup>8</sup>

If we sharpen the notion of 'leading to', this can explain the absence of contrafactives:

→ The fact of  $\neg p$  doesn't in any way 'lead to' the belief that  $p$ , nor is it precondition of belief that  $p$

We will formalize this notion this using **causal models**, drawing inspiration from Baglini & Francez (2016), Copley & Mari (2022) and Nadathur (accepted), a.o.

### 3.1 Causal models and impossible presuppositions

We adopt a simplified version of the causal model from Pearl (2000). A CM is a triple  $\langle U, V, E \rangle$ :

- $U$  is the set of exogenous variables (valued model-externally)
- $V$  is the set of endogenous variables (valued model-internally)
- $E$  is the set of structural equations  $f_i$  such that  $\forall v_i \in V, v_i = f_i(pa_i)$ , where  $pa_i \subset U \cup V$ <sup>9</sup>

The role of  $E$  is to tell us how to compute the values of each variable in  $V$  as a function of its parents ( $pa$ ).

These functions define dependencies in a causal model, which must additionally be representable with a directed acyclic graph. (**Note:** These arrows are not material implication!)

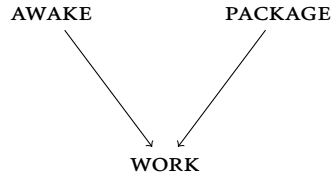
---

<sup>7</sup>Some theories assume that presuppositions are not lexically encoded, but derived pragmatically (e.g., Stalnaker 1970, 1973, 1974; Karttunen & Peters 1979; Kadmon 2001; Abusch 2002, 2010; Simons et al. 2010, 2016; Abrusán 2011, 2016). We believe that our account can be adjusted to fit this view by discussing relations between lexical entailments, rather than between presuppositions vs. at-issue content, but we leave a fully fleshed-out version for future work.

<sup>8</sup>See Schlenker (2021) for the 'epistemological precondition' view.

<sup>9</sup>We are considering idealized causal models here, without disturbances (in Pearl's sense).

For example:



**Context:** Whether you're at **WORK** in the morning depends on whether (a) you **WAKE** up and (b) you're expecting a **PACKAGE**.

**Exogenous variables:** { **AWAKE**, **PACKAGE** }

**Endogenous variable:** { **WORK** }

**Structural equation:**  $\text{WORK} := \text{AWAKE} \wedge \neg \text{PACKAGE}$

**Application:** If you're awake but expecting a package, you don't make it to work.

$\text{AWAKE} = \text{PACKAGE} = 1$

$\text{WORK} = 1 \wedge \neg 1 = 0$

In this model, the value of **WORK** is dependent on **AWAKE** and **PACKAGE**—there is a *causal chain* from each of **AWAKE** and **PACKAGE** to **WORK**.

Our proposal is that lexicalization is sensitive to this kind of chain, requiring a causal chain from presuppositions to at-issue content.

As a first pass:

(15) **Lexicalization constraint (preliminary, to be revised)**

A verbal predicate with at-issue content  $\alpha$  can have a presupposition  $\pi$  iff we can construct a causal model with a chain in which  $\pi$  is a **necessary precondition** of  $\alpha$ .

### 3.2 Assumptions: Causal relations involved in belief formation

Two sets of assumptions could link the presupposition and assertion of *know* and *shknow*:

One set concerns the relationship between truth of  $p$ , and evidence for  $p$  and  $\neg p$ .

The second concerns the relationship between evidence, acquaintance with evidence and belief.

#### Set of assumptions I: Truth and indicators of truth

Truths  $p$  are related to things in the world that indicate them, their indicators, in a constrained way.

(16) **Definition of indicator**

$i_p$  is an indicator of  $p$  iff  $i_p$  allows one to conclude that  $p$

1. The truth of  $p$  causes the existence of some **indicator** of  $p$

Example:

Because Konstanz is in Germany ( $p$ ), there are maps showing Konstanz in Germany ( $\text{indic}(p)$ ).

$$p \longrightarrow \text{indic}(p)$$

$$\text{if } p = 1 \text{ then } \text{indic}(p) = 1$$

2. The truth of  $p$  causes the absence of indicators of  $\neg p$ .<sup>10</sup>

<sup>10</sup>For brevity, this is deliberately too strong and should rather read: The truth of  $p$  doesn't cause the existence of indicators of  $\neg p$ . There may be maps showing Konstanz in France ( $\text{indic}(\neg p)$ ), independent of it being in Germany ( $p$ ). This shouldn't affect our argument. We discuss a perverse case below, the spy context, where  $p$  seems to cause  $\text{indic}(\neg p)$ .

Example:

Because Konstanz is in Germany, there are no maps showing Konstanz in France.

$$p \longrightarrow \text{indic}(\neg p)$$

$$\text{if } p = 1 \text{ then } \text{indic}(\neg p) = 0$$

The first assumption is used for factive reports, the second one, to block contrafactive reports.

### Set of assumptions II: Indicators and belief

We assume that  $Bp$  necessitates acquaintance with evidence that  $p$ .

1. An agent's acquaintance with any indicator for  $p$  causes belief that  $p$ :<sup>11</sup>

$$\text{acq}(i_p)(a) \longrightarrow B_ap$$

$$B_ap := 1 \text{ only if } \text{acq}(i_p)(a) = 1$$

2. Acquaintance requires existence:

Agents can become acquainted with an indicator only if it exists.

$$\text{indic}(p) \longrightarrow \text{acq}(i_p)(a)$$

$$\text{acq}(i_p)(a) := 1 \text{ only if } \text{indic}(p) = 1$$

### 3.3 Why we can know

Given our assumptions, because it is generally possible to construct causal models of the form below, where  $p$  and  $Bp$  are linked:

$$p \longrightarrow \text{indic}(p) \longrightarrow \text{acq}(i_p)(a) \longrightarrow B_ap$$

This means our constraint predicts predicates which presuppose  $p$  and assert  $Bp$  (*know*)<sup>12</sup> to be lexicalizable, since we can draw a chain fitting the above template.

- (17) [Context: Alice bought a present for Maxine (BUY) and hid it in the closet (CLOSET). Maxine saw the present (SEE) in the closet and formed the belief that it was for her (BELIEF).]

✓ Maxine knows that Alice bought her a present.

We can represent this with the following causal model, where  $p$  (BUY) is linked to  $Bp$  (BELIEF):

$$\text{BUY} \longrightarrow \text{CLOSET} \longrightarrow \text{SEE} \longrightarrow \text{BELIEF}$$

<sup>11</sup>We are referring not to belief writ large, but rather to something weaker: our intuitive folk concept of belief. So the relevant notion of 'evidence' here is deliberately broad, and could, for example, encompass things like religious faith.

<sup>12</sup>We are inspired by Goldman's (1967) causal theory of knowing, but our goal is not to provide a full analysis of knowledge reports. See Ichikawa & Steup (2018) for a survey.

- BUY makes the indicator CLOSET true.
- BELIEF requires acquaintance SEE and the indicator CLOSET to be true.

Two important notes:

- It doesn't suffice for BUY to be true for BELIEF to be true.  
In a different context, Maxine could have not seen the present.  
⇒ Makes knowledge ascription false.
- The link between BUY and CLOSET could be absent.  
In a different context, the present in the closet could be someone else's.  
⇒ Makes knowledge ascription #.

### 3.4 Why we can't *shknow*

False beliefs tend to be formed on the basis of erroneous information or unjustified assumptions:

- (18) [Context: Mary is looking at a map from 1980, but mistakenly thinks it is from 2023.]  
Mary shknows that Konstanz is in West Germany.

Intuitively, what makes *shknowledge* reports impossible is the fact of  $\neg p$  cannot lead to the belief that  $p$  in our causal sense.

Assume for the sake of proof by contradiction that the sentence in (19) is defined and true.

- (19) Marianne shknows that the Earth is flat.  
 $\rightsquigarrow$  The Earth is round (i.e., not flat).  
 $\rightarrow$  Marianne believes that the Earth is flat.

Because of the lexicalization constraint, there must then be a causal link between the fact that the Earth is round, and Marianne's belief that it is flat. What should that link be?

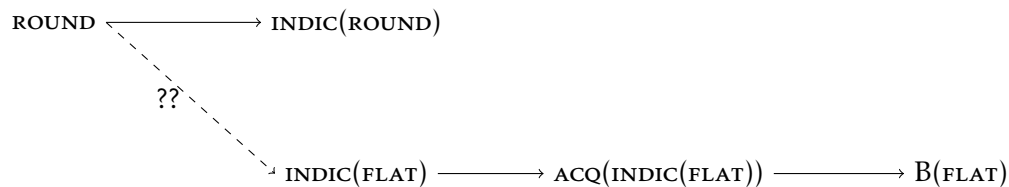
- By assumption, Marianne's belief is based on acquaintance with an indicator that the Earth is flat.

$$B(\text{FLAT}) = 1 \Rightarrow \text{ACQ}(\text{INDIC}(\text{FLAT})) = \text{INDIC}(\text{FLAT}) = 1$$

- Further assume that the fact that the Earth is round *does not generate evidence to the effect that it is flat*.

Then ROUND either inhibits or is unrelated to INDIC(FLAT) or is unrelated to INDIC(FLAT)

ROUND = 1  $\Rightarrow$  INDIC(FLAT) = 0 or ROUND and INDIC(FLAT) have no relation





It follows that either  $B(\text{FLAT})$  and  $\text{ROUND}$  are not linked, or that they impose a contradictory requirement on the model  $\Rightarrow$  we can't causally link  $Bp$  to  $\neg p$ .

**But!** Our lexicalization constraint is weak, so this isn't enough. It asks us to find any causal model from presupposition to at-issue content. We can do this for some contrafactive contexts!<sup>13</sup>:

- (20) [Context: Marianne thinks that Sasha is not a spy, because she is not behaving suspiciously. In reality, Sasha is a spy who is very good at concealing her spy-ness by acting normal.]  
Marianne *shknows* that Sasha is not a spy.

We can draw a causal fragment for this situation, capturing that spies act like non-spies to avoid arousing suspicion:

$$\text{spy}(s) \longrightarrow \text{indic}(\neg \text{spy}(s)) \longrightarrow \text{acq}(i_{\neg \text{spy}})(m) \longrightarrow B_m(\neg \text{spy}(s))$$

The issue is that Sasha's non-spy behavior underdetermines whether she is a spy or not: she would act that way no matter what, so her behavior can be construed as an indicator for  $s$  or  $\neg s$ .

The fact that we *can* construct models in which  $p$  causes  $\text{indic}(\neg p)$  means our constraint incorrectly predicts *shknow* to be a possible lexical item.

**Intuition:** The spy situation is 'abnormal'—the absence of *shknow* arises because propositions don't *generally* provide evidence for their negations, even if in some deviant cases they do.

What we need is a constraint which holds for *arbitrary* causal models of the relevant type.

In other words, the lexicalization constraint is sensitive to causal models of the form below, for arbitrary  $p$  (and attitude holder  $x$ ):

$$p \xrightarrow{\quad ?? \quad} \text{indic}(\neg p) \longrightarrow \text{acq}(i_{\neg p})(x) \longrightarrow B_x(\neg p)$$

We propose that these general structures are what attitude lexicalization is sensitive to:

- (21) **Lexicalization constraint, final:**  
A verbal predicate with at-issue content  $\alpha$  can have a presupposition  $\pi$  iff for every particular instance  $P$  of  $\pi$  and  $A$  of  $\alpha$ , if  $P$  and  $A$  are true, then we can construct a causal model with a chain in which  $P$  is a necessary precondition of  $A$ .

In other words, we need to be able to construct causal chains between  $p$  and  $B(\neg p)$  **for any choice of  $p$**  in order to lexicalize *shknow*

Because in the general case  $p$  is not linked with  $B(\neg p)$ , **contrafactive don't obey the constraint.**

## 4 Beyond (*sh*)*know*: Predictions

Our proposal predicts all attested presuppositional attitudes to adhere to the constraint. We believe this is so, illustrated with a couple of key examples.

<sup>13</sup>A more natural example comes from creatures that camouflage themselves—stick bugs, chameleons, and the like.

## 4.1 Attested predicates

**Response-stance predicates** such that *a verbs p*:

- presupposes that someone expressed a public commitment to *p*,
- asserts that *a* expresses *p* in response to that public commitment.

Cattell (1978): *accept, admit, agree, confirm, deny, reply, verify...*

- (22) a. Marianne accepted that the Earth was flat.  
 b. Marianne didn't accept that the Earth was flat.  
 $\rightsquigarrow$  Someone said that the Earth was flat.

Is a stimulatory causal link between *someone commits to p* and *a responds to p* generally possible? We believe yes. A toy model for *respond*:

$$\begin{aligned} committed - to(x)(p) &\longrightarrow respond(a)(p) \\ respond(a)(p) &:= 1 \text{ if } committed - to(x)(p) \end{aligned}$$

Intuition: *responding* describes acting towards a proposition in light of **someone's commitment**.

**Be wrong** seems to presuppose an attitude holder's commitment to *p* and assert the falsity of *p* (Anand & Hacquard 2014)

$$(23) \quad \llbracket \text{be wrong}_{A\&H} \rrbracket = \lambda p_{st} \lambda x_e : \mathbf{committed-to}(x)(p). \neg p$$

This doesn't adhere to our constraint on lexicalization! The subject's commitment to *p* obviously does not cause the fact of  $\neg p$

We propose to look at *wrong* from a different perspective: its at-issue content is not about the falsity of *p*, but rather the falsity of the presupposed **commitment**

In other words, we can treat *be wrong* is a response-stance predicate  $\rightsquigarrow$  conforms to the lexicalization constraint!

$$(24) \quad \llbracket \text{be wrong} \rrbracket = \lambda p_{st} \lambda x_e : \mathbf{committed-to}(x)(p). \neg \iota q [\mathbf{committed-to}(x)(q)]$$

This is not a big leap: it is truth-conditionally equivalent to (23), but differs in what entailments must be derived (vs lexically encoded):

- 'x is committed to *p*' and '*p* is false' entail that x's commitment is false
- 'x is committed to *p*' and 'x's commitment is false' entail that *p* is false

Admittedly, the difference between these two hypotheses for *wrong* is difficult to test empirically.

Some evidence suggests that *wrong's* at-issue content is not merely that *p* is false: *Wrong*-reports are natural if what is under discussion is the validity of a claim, rather than of *p* itself:

- (25) A: How were John's Oscar predictions?

- B: Awful. He was wrong that *Tár* would win Best Picture.
- (26) A: Did *Tár* get its richly deserved Best Picture Oscar?
- B: ?No, John was wrong that it would win.

## 4.2 Unattested predicates

Our lexicalization constraint also correctly rules out some predicates which are logically/conceivably possible but unattested (as far as we know).<sup>14</sup>

**Mirror *know***: Presupposes  $Bp$ , with at-issue content  $p$

- One's beliefs may be influenced by the facts of the world, but the truth of  $p$  is not influenced by our beliefs vis-à-vis  $p$  per se.
- Also holds of mirror versions of other factives (*regret*, *surprise*, etc.)

**Mirror response-stance predicates**: Presuppose (e.g.)  $x$  responded that  $p$ , with at-issue content *someone committed to*  $p$

- Responses have whatever they are responding to as a precondition

★ Causal chains of attested presuppositional predicates are unidirectional, so we expect mirror predicates not to exist!

## 5 Conclusion

We have provided additional support for the hypothesis that contrafactive predicates are not lexicalized in natural languages.

Our explanation for this gap in terms of a general constraint a lexical item's presuppositional content must causally implicate its at-issue content.

Attested predicates like *know* satisfy this constraint, but some unattested predicates do not.

At least two big questions remain.

### 1. What is the role of learnability?

One idea is that contrafactuals might not be lexicalized because they are (too) difficult to learn. But it is not impossible:

- Strohmaier & Wimmer (2022) find in an ANN experiment that contrafactuals are harder to learn than factives, but also easier to learn than non-factives like *believe* (!).
- Maldonado et al. (2022) find suggestive evidence that two predicates FALSEBEL ( $Bp \wedge \neg p$ ) and KNOWFALSE ( $\neg Bp \wedge \neg p$ ) are hard (but not impossible) to learn in experimental contexts.

<sup>14</sup>Predicates could also be unlexicalizable for other reasons. For example, we have no predicate which means 'know that the complement is false', even though it conforms to our constraint, but this is arguably reducible to a much more general fact: verbs don't lexicalize relations to the negation of their arguments (see Graff & Hartman 2011).

These findings are compatible with our analysis—perhaps the presupposition/at-issue relation we discuss is at the source of learnability difficulties.

## 2. What about presuppositions more generally?

It remains to be worked out how our analysis fits into the landscape of presuppositions and not-at-issue content more generally.

For instance, our account does not capture the behavior of many other presuppositions:

- Additive presuppositions (e.g. *too*)
- Gender presuppositions of pronouns
- Uniqueness presupposition of definite determiners

Many non-verbal triggers contribute *only* presuppositional content, so perhaps our lexicalization constraint is not relevant?

## References

- Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics and Philosophy* 34(6). doi:10.1007/s10988-012-9108-y.
- Abrusán, Márta. 2016. Presupposition cancellation: Explaining the ‘soft-hard’ trigger distinction. *Natural Language Semantics* 24(2). 165–202. doi:10.1007/s11050-016-9122-7.
- Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. In Brendan Jackson (ed.), *Proceedings of Semantics and Linguistic Theory (SALT)*, vol. 12, 1–19.
- Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27. 37–80.
- Albayrak, Samet. 2019. Semantic/pragmatic processing in Turkish propositional attitude verbs: The case of “zannet”. MA thesis, Middle East Technical University.
- Anand, Pranav & Valentine Hacquard. 2014. Factivity, belief and discourse. In Luka Crnić & Uli Sauerland (eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, vol. 1, 69–90. Cambridge, MA: MIT Working Papers in Linguistics.
- Anvari, Amir, Mora Maldonado & Andrés Soria Ruiz. 2019. The puzzle of reflexive belief construction. In M. Teresa Espinal et al. (ed.), *Proceedings of Sinn und Bedeutung* 23, 57–74.
- Baglini, Rebekah & Itamar Francez. 2016. The implications of managing. *Journal of Semantics* 33(3). 541–560.
- Bernard, Timothée & Lucas Champollion. 2018. Negative events in compositional semantics. In *Semantics and linguistic theory*, vol. 28, 512–532.
- Bledin, Justin. 2022. Composing negative individuals in truthmaker semantics. Amsterdam Colloquium presentation.
- Bossi, Madeleine. 2023. Negative bias, reminding, and pragmatic reasoning in Kipsigis belief reports. Ms., University of California, Berkeley.

- Cattell, Raymond. 1978. On the source of interrogative adverbs. *Language* 54. 61–77.
- Chemla, Emmanuel. 2008. An epistemic step for anti-presuppositions. *Journal of Semantics* 25(2). 141–173.
- Copley, Bridget & Alda Mari. 2022. *Forbid* is not *order not*. In *Proceedings of nels* 52, .
- Enguehard, Emile & Benjamin Spector. 2021. Explaining gaps in the logical lexicon of natural languages: A decision-theoretic perspective on the square of Aristotle. *Semantics and Pragmatics* 14.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung (On Sense and Reference). *Zeitschrift für Philosophie und philosophische Kritik* 100. 25–50.
- Glass, Lelia. 2022. The negatively biased Mandarin belief verb *yǐwéi*. *Studia Linguistica* 1–46.
- Goldman, Alvin I. 1967. A causal theory of knowing. *Journal of Philosophy* 64(12). 357–372.
- Graff, Peter & Jeremy Hartman. 2011. Impossible Predicates. GLOW 34 abstract.
- Hacquard, Valentine & Jeff Lidz. 2022. On the acquisition of attitude verbs. *Annual Review of Linguistics* 8(1). 193–212. doi:10.1146/annurev-linguistics-032521-053009.
- Hoeksema, Jack. 2021. Verbs of deception, point of view and polarity. In *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar*, 26–46. doi: 10.21248/hpsg.2021.2.
- Holton, Richard. 2017. Facts, factives, and contrafactuals. *Aristotelian Society Supplementary Volume* 91(1).
- Hsiao, Pei-Yi Katherine. 2017. On counterfactual attitudes: a case study of Taiwanese Southern Min. *Lingua Sinica* 3(4).
- Ichikawa, Jonathan Jenkins & Matthias Steup. 2018. The Analysis of Knowledge. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Metaphysics Research Lab, Stanford University Summer 2018 edn.
- Kadmon, Nirit. 2001. *Formal Pragmatics*. Malden, MA: Blackwell.
- Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyu Oh & David Dineen (eds.), *Syntax and Semantics 11: Presupposition*, 1–56. New York: Academic Press.
- Kierstead, Gregory. 2013. Shifted indexicals and conventional implicature: Tagalog *akala* ‘falsely believe’. Slides from SALT 23.
- Lohmann, Heidemarie & Michael Tomasello. 2003. The role of language in the development of false belief understanding: A training study. *Child Development* 74(4). 1130–1144. doi: 10.1111/1467-8624.00597.
- Maldonado, Mora, Jennifer Culbertson & Wataru Uegaki. 2022. Learnability and constraints on the semantics of clause-embedding predicates. In *Proceedings of the Cognitive Science Society annual meeting (CogSci) 2022*, .
- Nadathur, Prerna. accepted. Causal semantics for implicative verbs. *Journal of Semantics* .

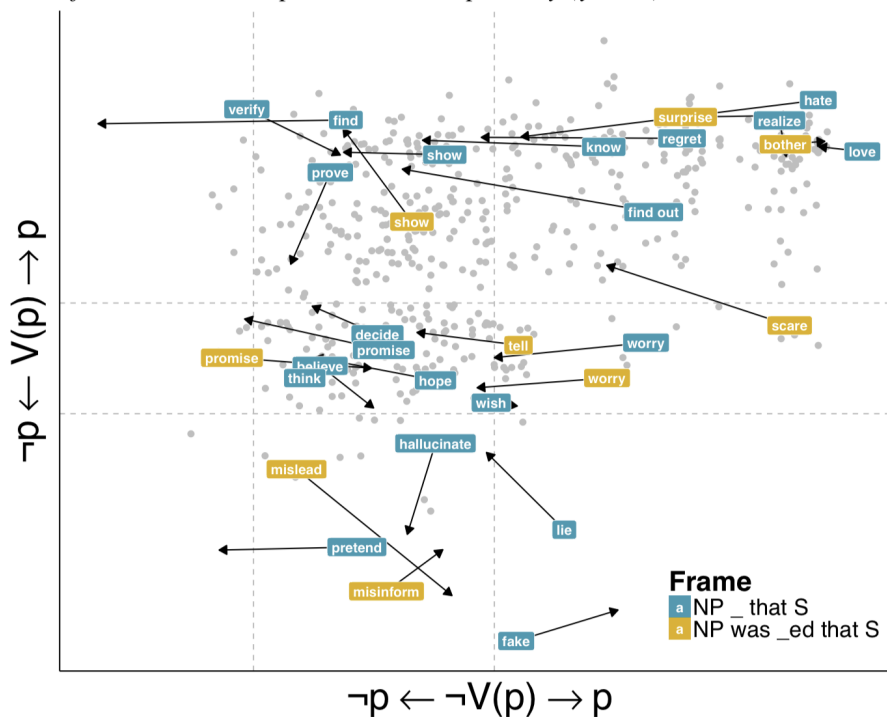
- Pearl, Judea. 2000. *Causality: Models, reasoning, and inference*. Cambridge University Press.
- Percus, Orin. 2006. Antipresuppositions. In A. Ueyama (ed.), *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science*, Japan Society for the promotion of science.
- Phillips, Jonathan & Aaron Norby. 2021. Factive theory of mind. *Mind & Language* 36. 3–26. doi:10.1111/mila.12267.
- Roberts, Craige & Mandy Simons. 2022. Preconditions, presuppositions, and projection. In Marco Degano et al. (ed.), *Proceedings of the 23rd Amsterdam Colloquium*, 429–436.
- Roberts, Thomas. 2021. *How to make believe: Inquisitiveness, veridicality, and evidentiality in belief reports*: University of California, Santa Cruz dissertation.
- Roberts, Tom. 2019. I can't believe it's not lexical: Deriving distributed veridicality. In Katherine Blake, Forrest Davis, Kaelyn Lamp & Joseph Rhyne (eds.), *The Proceedings of Semantics and Linguistic Theory (SALT) 29*, 665–685.
- Sander, Thorsten. ms. The K's evil twin. Unpublished manuscript, Uni Duisburg.
- Sauerland, Uli. 2007. Implicated presuppositions. In Anita Steube (ed.), *Sentence and context*, Mouton de Gruyter, Berlin, Germany.
- Schlenker, Philippe. 2021. Triggering Presuppositions. *Glossa: a journal of general linguistics* 6(1).
- Simons, Mandy, David Beaver, Craige Roberts & Judith Tonhauser. 2016. The best question: Explaining the projection behavior of factives. *Discourse Processes* 54(3). 187–206. doi: 10.1080/0163853X.2016.1150660.
- Simons, Mandy, Judith Tonhauser, David Beaver & Craige Roberts. 2010. What projects and why. In N. Li & D. Lutz (eds.), *Proceedings of SALT 20*, 309–327. CLC.
- Stalnaker, Robert. 1970. Pragmatics. *Synthese* 272–289. doi:10.1007/BF00413603.
- Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic* 2(4). 447–457. doi: 10.1007/BF00262951.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton K. Munitz & Peter Unger (eds.), *Semantics and Philosophy*, 197–213. New York: NYU Press.
- Strohmaier, David & Simon Wimmer. 2022. Contrafactuals and learnability. In Marco Degano et al. (ed.), *Proceedings of the Amsterdam Colloquium 2022*, 298–305.
- de Villiers, Jill G. & Jennie E. Pyers. 2002. Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief-understanding Author links open overlay panel. *Cognitive Development* 17(1). 1037–1060. doi:10.1016/S0885-2014(02)00073-4.
- White, Aaron Steven & Kyle Rawlins. 2018. The role of veridicality and factivity in clause selection. In Sherry Hucklebridge & Max Nelson (eds.), *Proceedings of the 58th Annual Meeting of the North East Linguistic Society*, 221–234. Amherst, MA: GLSA Publications.
- Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford University Press.

## A No contrafactuals in the lexicon of English

The following graph from White & Rawlins (2018) plots attitude predicates in the MegaVeridicality dataset (1000 English clause-embedding predicates) in terms of how factive (x-axis) and veridical (y-axis) they were judged to be.

Of interest is the relative emptiness of the bottom two quadrants (anti-veridicals), and of the bottom left quadrant in particular (contrafactuals).

- (14) *Normalized responses for contexts with negative matrix polarity (x-axis) against those for contexts with positive matrix polarity (y-axis)*



## B Are contrafactuals just not utilitarian?

An alternative to our causal approach: Lexica as a whole optimize an informativity-cost tradeoff; contrafactual verbs do not exist because they are too costly and/or uninformative

In this view, lexical gaps arise when the cost of adding additional lexical items outweighs their informative usefulness.

One way to be uninformative is to describe situations that are either statistically marginal:

- There is no lexeme for *tapdancing in the manner of a purple giraffe*, because its use cases are (unfortunately) vanishingly rare

This idea has been profitably used to analyze other kinds of lexical gaps, especially in logical vocabulary, e.g. in the quantifier domain (Enguehard & Spector 2021).

We like the intuition that contrafactuals are not utilitarian enough to lexicalize, but frequency alone seems not to be the whole story.

First, *prima facie*, false belief situations don't seem inherently outlandish:

- (27) CG: Deniz isn't French.
- a. Many people think that I'm French. (And ask me judgments.)
  - b. Many people shknow that I'm French. (And ask me judgments.)
- (28) CG: Deniz isn't French.
- a. My friends don't think that I'm French. (And they don't ask me for judgments.)
  - b. My friends don't shknow that I'm French. (And they don't ask me for judgments.)

Second, if utility of a verb is calculated in relation to its frequency of possible uses, is not clear how we could quantify this.

- We might think that speaker-subject belief misalignment is more marked (and therefore rarer) than alignment:
  - False belief reports often misunderstood by preschool children (de Villiers & Pyers 2002; Lohmann & Tomasello 2003, a.m.o.; see also citations in Hacquard & Lidz 2022)
  - False beliefs are arguably representationally complex, since agents cannot simply map their model of the world onto another's mental state (Phillips & Norby 2021)
- Unlike e.g. quantifiers, *believe* and *(sh)know*-reports may also compete against non-substitutive alternatives like bare assertions, at least in 1sg-subject contexts

A utility-based approach may well provide a suitable explanation in the end for the contrafactive gap, but we we leave operationalizing this hypothesis to future work.